WPS4534

Policy Research Working Paper 4534

The Extent of the Market and Stages of Agricultural Specialization

M. Shahe Emran
Forhad Shilpi

The World Bank
Development Research Group
Sustainable Rural and Urban Development Team
February 2008

Abstract

This paper provides empirical evidence of nonlinearity in the relationship between crop specialization in a village economy and the extent of the market (size of the urban market) relevant for the village. The results suggest that the portfolio of crops in a village economy becomes more diversified initially as the extent of the market increases. However, after the market size reaches a threshold, the production structure becomes specialized again. This evidence on the stages of agricultural diversification is consistent with the stages of diversification identified in the recent literature for the economy as a whole and also for the manufacturing sector. The evidence highlights the importance of improving farmers' access to markets through investment in transport infrastructure and removal of barriers to trading.

This paper—a product of the Sustainable Rural and Urban Development Team, Development Research Group—is part of a larger effort in the department to understand the impact urban market size on agricultural commercialization and development. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at fshilpi@worldbank.org.

The Extent of the Market and Stages of Agricultural Specialization

M. Shahe Emran[1]
Department of Economics and ESIA
George Washington University

Forhad Shilpi
The World Bank

JEL Codes: O12, R11, R12

Key Words: Structural Change, Agriculture, Crop Specialization, The Extent of the Market, Market versus Home Production, Commercialization

---

[1]Emails for correspondence: shahe.emran@gmail.com (M. Shahe Emran), fshilpi@worldbank.org (Forhad Shilpi). We would like to thank Bryan Boulier, Tony Yezer and Stephen Smith for helpful discussions. The standard disclaimers apply.

1 Introduction

Understanding the process of structural change has been a central focus of development economics from Lewis (1954), Kuznets (1973), Chenery et. al. (1986) to Lucas (1988, 2004), among others. Most of the theoretical and empirical literature on structural change and long-run evolution of an economy focuses on the transition from a predominantly agrarian and rural economy to an industrialized and urban one (Chenery et. al., 1986; Locay, 1990 ; Laitner, 2000 ; Lucas, 1988; Buera and Kaboski, 2006; Matsuyama,2005). For a large number of developing countries where agriculture still predominates the economic landscape, the issue of structural transformation within agriculture– from a traditional subsistence based agriculture to more specialized and market oriented one– is, however, equally important. This is because structural transformation of an economy into more diversified non-agricultural (non-farm and industrial) activities is frequently triggered by productivity growth and increasing commercialization and specialization in agriculture (Johnson, 2000; Gollin, Parente and Rogerson, 2002, 2006). This paper presents an empirical analysis of structural change within agriculture with a focus on the role played by the extent of the market.

The idea that the extent of market is a principal driving force behind specialization dates back at least to Adam Smith (Smith, 1776).[1] In this classic Smithian account, a larger market allows greater division of labor and specialization by ensuring adequate demand for specialized skills and products. The more recent literature has underscored the importance of a large market in the adoption of increasing returns technologies that facilitates greater specialization in intermediate inputs and leads to higher economic growth (Murphy, Shleifer, and Vishny, 1989; Rodriguez-Clare, 1996; Ciccone and Matsuyama, 1996).[2] An implication of this literature is that there is a monotonic relationship between the extent of the market and the degree of specialization. The recent literature on the 'stages of diversification',

[1] "As it is the power of exchanging that gives occasion to the division of labor, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market" (Smith, A, 1776, Book I, Chapter III). See also Stigler (1951).

[2] The positive influence of market size on specialization implies a positive correlation between initial income and subsequent growth of a country. Ades and Glaeser (1999), using cross country growth regressions, finds strong positive correlation between initial income and subsequent economic growth particularly for relatively closed economies. These results are further confirmed by Alcala and Ciccone(2003).

however, uncovers non-linearity in the process of specialization. According to this literature, the production structure initially becomes more diversified as per capita income grows; and only after a threshold level of income is reached, the production structure becomes more specialized (Imbs and Wacziarg, 2003, Kalemli-Ozcan et. al. 2003). This inverted U pattern in stages of diversification holds both in the aggregate economy and within the manufacturing sector. This literature, however, does not address the pattern of structural change within the agricultural sector.

In the context of agriculture, a farmer's decision to grow certain crop(s) and to participate in the market is governed, among other things, by her/his perceived price and yield risks and subsistence considerations ("survival first"). Apart from its role in facilitating greater division of labor a la Adam Smith, the extent of market has important implications for the price risk faced by the farmers both in input (e.g. fertilizer and pesticide) and output markets. When the relevant urban market is small, the price risk is likely to be high due to imperfect matching in a thin market. This induces farmers to stick with subsistence farming in an attempt to avoid starvation. An increase in the extent of the market leads to higher and relatively less volatile price for non-staple crops. This might induce the farmers to allocate some land to non-subsistence crops, but the subsistence considerations along with price and yield risks in an environment where insurance markets are missing force farmers to adopt a more diversified crop portfolio rather than complete specialization in non-subsistence crops. When the extent of market reaches a threshold, the price risk is reduced significantly due to better matching. Proximity to a large urban center may also open up possibilities for better risk bearing capacity through access to formal credit and insurance markets. Moreover, a large urban market allows scale economies in both production and marketing of non-cereal crops like fruits and vegetables. The twin forces of lower price risk and increasing returns in production and marketing (transport and storage) when strong enough can induce the farmers to start specializing in the non-subsistence crops production for the urban market. The interplay of subsistence and risk considerations and the scale economies thus implies that the relationship between agricultural specialization and the extent of market is likely to be non-linear.[3]

We use data from the Nepal Living Standard Survey (NLSS) of 1995/96 to uncover the role played by the extent of the market in agricultural specialization and commercialization at the village level. Crop agriculture in Nepal, like many other developing economies, is characterized by low degree of commercialization and specialization on average. There are, however, striking differences among villages in terms of the level of agricultural development covering the entire range from completely specialized production of non-staple crops to nearly complete subsistence agriculture. The stark geographical differences in Nepal also resulted in large variation in sizes of the urban centers (from population of 10,000 in smaller towns to 421,000 in the capital city Kathmandu in 1991). These large variations in the level of agricultural development along with that in access to and size of urban markets enable us to empirically characterize the relationship between agricultural specialization and the extent of the market. The results based on the Nepalese data are, however, of more general interest as they are likely to be applicable to many other developing countries which are characterized by relative isolation of rural areas due to poor infrastructure as well as low level of agricultural development.[4]

We analyze two dimensions of structural change: the pattern of product diversification (crop specialization), and the degree of market production as opposed to home production (i.e., commercialization of agriculture). The pattern of crop specialization is measured by Herfindahl index of concentration of cropland use, and by the share of land devoted to non-cereal crops. Sales of non-rice crops and all crops as a percentage of total production are taken as measures of commercialization. Empirical estimation of this paper introduces several improvements. First, as opposed to the standard practice of defining market as the nearest urban center, we allow for the possibility that villagers may be trading at multiple urban centers. We define the extent of the market as an average of the gravity measures of all the markets where villagers may trade. The gravity measure itself is defined as a ratio of urban

[3] For an interesting theoretical analysis that highlights the tension between the increasing returns and risk in determining technological change and production structure in agriculture, see Dixit, 1993.

[4] These would include regions like Northern India and Pakistan, and many other developing countries in Africa and Asia.

market size measured by its income level and some non-linear transformation of the distance to that market center. Second, the extent of the market is instrumented using geographic characteristics to allay any concerns regarding potential reverse causation. Finally, in addition to parametric regression analysis, this paper utilizes semi-parametric techniques to uncover the nonlinearity in the effects of the extent of market on agricultural specialization.

The parametric regression results show that the extent of the market has a statistically significant (at 1 percent level) positive effect on land allocated to non-cereal crops and on crop sales. In the case of Herfindahl index of concentration of cropland use, the effect of a larger market is negative, but lacks statistical significance (not significant at 5 percent level). The semi-parametric estimation indicates that the relationship between Herfindahl index and market size is U-shaped. In contrast, the relationship is monotonically upward sloping in the case of share of land allocated to non-cereal crops as a function of the market size. The relationship between sales of non-cereal (non-staple) crops and market size is monotonically increasing with no significant evidence of nonlinearity. These results imply that when the farmers have access only to small urban markets, the production structure in a village economy tends to be specialized in subsistence agriculture with most of the land devoted to a single subsistence crop (rice in case of Nepal) and only a limited degree of commercialization. As the extent (size) of the market increases, the portfolio of crops in a village economy becomes more diversified initially. However, after the market size reaches a threshold, the production structure starts to specialize again. This evidence on the stages of agricultural diversification thus mirrors the stages of diversification identified in the recent literature for the economy as a whole and also for the manufacturing sector (Imbs and Wacziarg, 2003). Our results also show that there is a monotonically increasing relationship between the degree of market participation (commercialization) among the households in a village and the extent of the relevant urban market. To the best of our knowledge, this paper provides the first evidence on the stages of agricultural specialization as it relates to the extent of the (urban) market[5].

[5] Although there is a large literature on the role of urban market in agricultural specialization, the focus of that literature is on the *access* to urban markets (see, for example, Jacoby, 2000; Fafchamps and Shilpi 2003, and 2005). For more general discussion on the role of urban markets in agricultural diversification and

The rest of the paper is organized as follows. Section 2 describes the simple conceptual framework underpinning the empirical analysis. Section 3 provides details about the data base used in the empirical analysis. Section 4, organized in a couple of subsections, presents the main empirical analysis and results. The paper is concluded in Section 5.

2 Conceptual Framework

To explore the relationship between market size and agricultural specialization, we start with the simple 'gravity model' which can be expressed as:

$$y_i^z = f^z(D_i, M_i, X_i) + u_{iz} \tag{1}$$

Where D_i is the distance to the urban market center relevant for village i, M_i is the size of that market, X_i is a vector of other relevant control variables and u_i is the error term. The vector y^z is a vector of dependent variables which in our case includes different measures of specialization and commercialization, to be defined precisely in the following paragraphs. In its simplest form, the gravity model is then estimated in a semi-log-linear specification as the following:

$$y_i^z = \beta_{z0} + \beta_{z1}\log(D_i) + \beta_{z2}\log(M_i) + X_i'\gamma_z + u_{zi} \tag{2}$$

The estimation of equation (2) requires identification of the most relevant market (s) for the households residing in village i. The standard practice in the current literature is to use the city/urban center closest to the village as the relevant market. However, this is simplistic and can be misleading as the empirical results reported later show. To see the importance of including more than one city/town as the relevant market, consider the case shown in Figure 1; city A is located nearest to village i, city B is only a short distance farther but of a much larger size. It is quite likely that compared with city A, urban demand in city B

specialization, see World Development Report, 1982; World Bank, 2006; Binswanger and Deininger, 1997.

exerts stronger influence on farm households' crop choice and market participation decisions. Moreover, the households in village i may trade different goods in different cities in which case distance to only the nearest city or market may not be the relevant measure of the extent of the market. The effective market for a village, given the transport infrastructure, consists of all the city/town market centers around a village where the villagers participate in. In order to allow for the possibility of trading by households of a village at multiple urban locations, we extend the traditional definition of gravity measure and define the effective market size for village i as:

$$M_i^e = \sum_{j=1}^{K} \omega_{ij} m_{ij} \tag{3}$$

where m_{ij} is the market size of urban center j and $1, 2, .., K$ are the urban centers where villagers trade, and ω_{ij} is the weight of urban center j.The weight for each market j is defined as:

$$\omega_{ij} = \frac{\delta(d_{ij})}{K}$$

where d_{ij} is the distance from village i to urban center j. We assume that $\frac{\partial \delta(.)}{\partial d_{ij}} < 0$ and $\frac{\partial^2 \delta(.)}{\partial d_{ij}^2} > 0$. These assumptions about the functional form of $\delta(.)$ imply that farther market centers are given smaller weights compared with the closer market centers. We assume that $\delta_{ij} = 1/d_{ij}^2$. Substituting $\delta(.)$ in equation (3), we have:

$$M_i^e = \frac{1}{K} \sum_{j=1}^{K} \frac{m_{ij}}{d_{ij}^2} = \frac{1}{K} \sum_{j=1}^{K} g_{ij} \tag{4}$$

where g_{ij} is the gravity measure widely utilized in studying the effect of market size on international trade flows. The effective market size for village i, as defined in equation (4), is thus an average of gravity measures of urban centers (1,2,...,K) where residents of village i trade. It is clear from equation (4) that closer and larger urban markets are given higher weights in defining the effective market size. Moreover, this definition of effective market size incorporates the negative and non-linear effect of distances on farm households' cropping pattern and market participation decisions as emphasized in recent empirical evidence on spatial organization of economic activities in rural areas (Jacoby, 2000; Fachamps and Shilpi, 2003). We also check robustness of our empirical results using alternative specification of the weight function. The empirical specification of M_i^e requires prior knowledge about K, the number of urban centers relevant for a village. As villagers may go to different markets for trading different products, empirical estimation is done for different values of K in order to establish robustness of our empirical results.

Using the semi-log formulation again, the estimating equation can be specified as:

$$y_i^z = \beta_{z3} + \beta_{z4} \log(M_i^e) + X_i'\gamma_z + u_{zi} \tag{5}$$

The linear specification in equation (5) may be inadequate to study the effect of the extent of market on agricultural specialization. The non-linearity in the relationship between agricultural specialization and market size may arise from a number of factors. First, in the case of crop agriculture, the yield and price risks are among the most important determinants of a farm household's land allocation across crops (Roumasset, 1976 , Newbery and Stiglitz, 1981, Islam and Thomas, 1996). When the relevant urban market is small, the price risk is likely to be higher due to imperfect matching in a thin market. This is especially important for non-staple (non-cereal) produce like fruits, vegetables and spices for which the extent of market is much more limited in a typical developing country because of the Engel's Law. This implies that a farmer facing a small urban market might not specialize in the production of high risk and potentially high return non-subsistence crops like fruits and vegetables although she might be willing to devote some land to such crops at the margin. When the extent of market reaches a threshold, the price risk is reduced significantly because of better matching in a thick market. A larger market also ensures adequate demand for large scale production and higher profit for non-staple crops. The higher profit may be due to more favorable prices for both inputs and output, and adoption of increasing returns technology and agglomeration effects.[6] A large urban market allows scale economies in marketing of non-cereal crops like

[6] Note that a larger market is likely to induce farmers to adopt new technology in the production of the subsistence crop like rice also. This helps in specialization and commercialization as the farmers can allocate more land to non-subsistence crops without worrying about its own subsistence requirements.

vegetables (fixed costs in transportation and storage by the wholesale traders) which in turn translates into better prices for farmers if entry into marketing of agricultural goods is not restricted. Another important point is that access to large urban areas means that the rural households have access to a rich set of markets including credit (banks) and insurance markets. As is well known, farmers are likely to behave in a risk averse fashion when markets are incomplete, especially when credit and insurance markets are missing (Newbery and Stiglitz, 1981). A more complete set of markets allows the farmers to take more production risk and devote more land to non-subsistence and cash crops. The interplay of subsistence and risk considerations, urban demand pattern and scale economies is likely to result in a non-linear relationship between agricultural specialization and the extent of the market.

Equation (5) can be modified to allow for flexible functional form with respect to market size as the following:

$$y_i^z = \beta_{z5} + g_z(M_{ji}^e) + X_i'\gamma_z + u_{zi} \tag{6}$$

where $g_z()$ is expected to be a non-linear function of M_{ji}^e.

However, the estimation of equation (5) can be implemented only when we have a measure of market size of each urban centers. Market size in a city is often represented by its population density (Ciccone and Alcala, 2003; Fafchamps and Shilpi, 2003 and 2005). This, however, does not take into account of the fact that the pattern of consumer demand in a city depends critically on the level of income of its population. According to the Engel curve relationship, poor people tend to spend a higher proportion of their income on staples relative to the non-poor. Moreover, with an increase in income, demand for non-staples rises more sharply than the demand for staples. Thus, it is the level of urban income which is likely to exert discriminating influence on demand for different agricultural crops and consequently on agricultural specialization. Despite the presence of a positive and significant correlation between income and population density within an urban center, population density as an indicator of urban demand is likely to be imperfect at best. We thus use the total urban income as a measure of the extent of the market.

As to the vector of dependent variables y_i^z in equation (4), three measures of agricultural specialization and commercialization are analyzed in this paper. First, we define a Herfindahl index of concentration of crop land use as:[7]

$$S_j = \sum_{h=1}^{H} (\frac{l_{jh}}{lj})^2 \tag{7}$$

where l_{jh} is the amount of land devoted to crop h in village j, l_j is the total amount of land farmed in village j and H is total number of crops grown. Notice that if all land in a village is devoted to one crop, then specialization index S_j is equal to unity. The more the number of crops grown in a village, the lower is the value of S_j. After controlling for crop suitability and land quality, a complete specialization ($S_j = 1$) could result from all land being devoted to non-cereal crops due to commercialization of agriculture. It could also be associated with subsistence nature of agriculture where virtually all land is allocated, for instance, to a cereal crop like rice in case of Nepal. This implies that the nature of specialization (whether due to commercialization or to subsistence nature of agricultural production) can be discerned only if the Herfindahl index of specialization is compared with an indicator of non-subsistence specialization. The share of land devoted to non-cereal crops at the village level is used as a measure of non-subsistence specialization. Our analysis focuses mainly on specialization and diversification during the dry season. This is due to the fact that the cropping pattern during the wet season is completely dominated by rice, a crop which benefits greatly from the monsoon conditions. Because of heavy monsoon and inundation associated with it, farmers have few options other than cultivating rice during the wet season. Sales of all crops and non-rice crops as a percentage of their respective production in the village are taken as measures of agricultural commercialization. It should be noted that the measures of specialization and commercialization are defined at the village level.

[7] This index is similar to the specialization index in Imbs and Wacziarg (2003) and occupational specialization index used in Ades and Glaeser(1999).

3 Data

The data for the empirical analysis of this paper come from the Nepal Living Standard Survey (NLSS) of 1995/96. The NLSS is a nationally representative survey which collected information from 3373 households spread among 274 primary sampling units (locally known as 'wards') covering 73 of Nepal's 75 districts. In addition to the comprehensive information on household and its members' characteristics, and household's expenditure levels, the survey collected detailed information on agricultural activities including cropping pattern, crop production and sales. Of all households for which we have complete information (about 3344), 75% are engaged in crop production. Nearly all of the farmers engaged in crop production are located in rural areas (93%) and the rest located in and around rural towns. The farm households (about 2531 households) are distributed in 257 wards/villages. We drop wards/villages with less than three farm households from our sample. The empirical analysis of this paper is thus based on the sample of 237 villages/wards where at least three households are engaged in agricultural production. Both dependent and explanatory variables for the empirical analysis are defined at the village level using the information on farmers residing in respective ward/village.

Panel A of Table 1 provides the summary statistics for different measures of agricultural specialization and commercialization. The Herfindahl index (HI), defined in equation (7), is constructed from the cropping pattern observed at the village level. As already noted, the index takes a value of unity if all land in a village is devoted to a single crop and declines in value with an increase in the number of crops grown in a village. The median of the Herfindahl index is about 0.27 (mean=0.31). There are, however, considerable variations in the level of specialization across villages covering the entire spectrum from complete specialization (HI=1) to highly diversified cropping patterns. Our second measure of specialization (non-subsistence specialization) is the share of land devoted to production of relatively high value non-cereal crops (fruits, vegetables, oilseeds, spices, cash and other crops). According to Table 1, cereal crops dominate the cropping pattern. On average, less than a third of total land during dry season is used to produce non-cereal crops, the median is smaller about 0.25.

However, there is considerable variation across villages, ranging from no land to all land allocated to non-cereal crops.[8]

Our final measure relates to commercialization of agriculture which captures the structural change away from home to market production emphasized in the recent literature (see, or example, Gollin et, al. 2002, 2006). Specifically, the shares of production of all crops and non-rice crops sold in the market are taken as measures of commercialization.[9] According to Table 1, on average, about 14.7% for rice output, and 21% of non-rice output are sold in the market. Overall, 19% of total crop output (median=15%) is sold for cash, rest being consumed at home or handed out as in kind payments. Despite the relatively low average degree of commercialization, in a number of villages, more than 50% of output are sold even in the case of rice, the main staple crop in Nepal.

Panel B of Table 1 reports the summary statistics for access to and size of the market located nearest to the surveyed villages. Despite its comprehensive data coverage, the NLSS 1995/96 lacks information on access to urban centers as well as market size in each urban centers. We complement the NLSS data by constructing measures of both access to urban centers and urban income. The Population Census of 1991 identifies 34 towns and cities in Nepal where a town is defined as a settlement of more than 10, 000 inhabitants. We first compute the distance between each surveyed ward/village and each of these towns. Distances are normally taken along existing roads, except when roads do not exist, in which case we calculate the shortest arc distance to the nearest road, and then the distance to various cities along the road. Distances are then converted into travel time using available information about trucking and walking speeds along various types of roads in Nepal.[10] Off the road

[8] Both the Herfindahl Index and share of land devoted to non-cereal crops have a maximum of 1. There is, however, only one village with H=1. There are only 5 villages where all of the land is allocated to non-cereal crops. This should ameliorate any concerns for censoring in our data.

[9] The regression analysis focuses on all crops and non-rice crops sales because there is no rice sales in a large number of villages (about a third of the villages). This causes the problem of censoring in the case of regression for rice sales. The difference between the rice and non-rice sales can be highlighted by focusing on all vs. non-rice sales as well while avoiding the censoring problem.

[10] Travel speeds are calculated for various terrains and types of road. The travel times on different terrains and road types were obtained through discussion with various transportation experts and South Asia operations staff at the World Bank. Travel on highways and provincial roads is assumed to take place by truck; travel on secondary roads is assumed to be by cart. For details on the construction of the travel times, see Fafchamps and Shilpi(2003).

travel is assumed to take place by foot – a reasonable assumption for Nepal given the nature of the terrain. The median distance from surveyed wards to nearest town is about 2 hours and 21 minutes. The mean distance is however much higher; about 4 hours and 26 minutes, because a number of villages are located far off from nearest towns (Figure 2a). Indeed, about 14% of villages are located at least 10 hours or more from the nearest town, the farthest one being about 29.5 hours away. Such wide variations in access to urban markets are outcomes of striking geographical disparities in Nepal, a country which extends from the relative plains of Terai to high Himalayas.

Data on city/town population are available from the Population Census of 1991. The population density of towns/cities in Nepal displays wide variations; the smallest town had barely 10,000 people residing in it. The largest city, Kathmandu, on the other hand had a population of 421,000 in 1991. We estimated per capita consumption expenditure for urban residents from the NLSS data. Total urban income is derived by multiplying per capita income by urban population. In the cases of smaller towns, per capita expenditure data are not always available. In those cases, average per capita consumption expenditure in the district where town is located is used to compute total urban income. The median income in the nearest town/city is about Rs. 302 million, and mean about Rs. 1.2 billion. Urban income displays high degree of variations (Figure 2b). Using the formula in equation (4) and assuming $K = 34$ (total number of towns/cities in Nepal in 1991), we define the effective market size M_i^{eA}.[11] The large variations in travel times and in urban income resulted in even larger variation in the measure of the extent of market at the village level (SD= 573 and mean=105). These variations help us explore to the relationships between different measures of agricultural specialization and the extent of the market.

[11] As we discuss later, our results are not sensitive to alternative definitions of market size corresponding to the different values of K chosen.

4 Empirical Results

We present the empirical results in a sequential manner starting with the simple gravity specification as in equation (2) and then proceeding to equation (5). To isolate the effects of the extent of the market, we control for a rich set of factors that may determine agricultural specialization at the village level including a wide variety of village level characteristics. The set of explanatory variables includes average household size and composition (share of adult female members, share of children, share of old etc.) in the village.[12] In the sales regressions, the household size and composition variables control for possible subsistence considerations whereas in the case of land allocated to non-cereal crops and of Herfindahl index of concentration of cropland use, they control for labor supply and gender specialization. The average education level of adult male and adult female in the village are also introduced as possible controls for average human capital in the village. The dependent variables in the sales regressions are already normalized by production levels. In addition, we include a number of farm characteristics that can influence farm productivity and hence sales. The average characteristics of owned land at the village level are used as explanatory variables instead of that of operated land. These characteristics include size of owned landholding, a number of characteristics of owned land including land quality (share of khet land which is especially suitable for rice production, share of irrigated land, share of land of different soil quality such as awal, dwaim or sim), number of farm animals (cows and buffaloes), and value of farm equipments. We also include dummies for different agro-ecological belts (Mountains, hills etc.). The average per capita consumption expenditure of the village is introduced to control for possible impact of local demand as well as level of local development. In order to control for access to credit, we included two regressors: whether there is a bank in the village, and the average remittance income of the households residing in the village.

The parametric regression equations (2) and (5) can be estimated using Ordinary Least Square (OLS).[13] The simple OLS regressions however ignore possible endogeneity of the

[12] The omitted category is share of adult male.

[13] In all of the regressions reported in this paper, the standard errors are corrected for heteroskedasticity.

market size; larger urban areas may be supported by better agricultural potentials of the surrounding areas. In addition, access to market is likely to be endogeneous as the placement of roads may be targeted to areas with better agricultural potential. To correct for the potential endogeneity problem, we rely on instrumental variable (IV) estimation. The regressors capturing the effect of the extent of market in equation (2) and (5) involve non-linear transformation of both urban income and travel time to urban areas. Because of this non-linearity, we follow an IV estimation procedure suggested by Woolridge(2007). At the preliminary stage of estimation, we estimate two 'instrumenting regressions'. The dependent variable in the first regression is the urban income of all Nepalese cities. The travel time to cities from all of the sampled wards/villages is the dependent variable in the second instrumenting regression.[14] The predicted values from these two regressions are used to define the relevant market size variables using the same formulas as those used in equation (2) and (5). The predicted market size variable is then used as the instrument for the actual variable in the IV estimation of equation (2) and (5). The advantage of this approach is that the usual standard errors and test statistics are asymptotically valid (Woolridge, 2007).

Following the recent literature, the identifying instruments in the preliminary regression for log of urban income are derived from physical/geographic characteristics of the district in which the city/town is located (see, for example, Frankel and Romer, 1999, Fafchamps and Shilpi, 2005). The set of instruments includes log of size of the district in square kilometers, log of its arable land area, log of the distance to nearest river, and the mean and standard deviation of elevation within the district. We also include as regressor a dummy if the district is in the mountaineous part of the country. The area of the district is included as bigger district may support larger population and hence higher total urban income. Arable land proxies for food production potential and distance to navigable part of the nearest river for ease of access. Elevation controls for climatic condition as towns are unlikely to be located in higher elevation. The mountain dummy as well as standard deviation of elevation control for roughness of the terrain. The regression result is reported in Appendix Table A.1. The R^2

[14] We emphasize here that these preliminary 'instrumenting regressions' are not the standard first stage regressions in 2SLS.

of the regression is 0.43. The strong multicollinearity of the regressors, however, led to none of the individual regressor being statistically significant (mean variance inflation factor=8.7) although they in general have expected sign.

Travel time between a village and a town is instrumented using foot travel time as well as physical characteristics of the village/ward and town. The foot travel time is computed using iso-elevation curves to account for the mountainous nature of the terrain. The regression result is reported in appendix Table A.2. As expected, foot travel time is by far the most important determinant of travel time. Other regressors are also significant and usually with expected sign. The instruments account for considerable variation in travel time with a R^2 equal to 0.84.

4.1 Preliminary Results

Table 2 reports the regression results based on our simplest specification implied by the standard gravity model outlined in equation (2). Also, the standard definition of relevant market as the total income of the nearest city/town is adopted. The upper panel reports the results when equation (2) is estimated by OLS. The market size variable is statistically significant with expected positive sign only in the regression for share of land allocated to non-cereal crops. The travel time to the nearest city/town has the expected negative sign in all regressions except for that of Herfindahl index. This access to market variable is also statistically significant in all regressions except for Herfindahl index.[15]

The results from IV regressions are presented at the lower panel of Table 2. The instruments used for travel time to nearest town/city and income of that town/city are predicted travel time to nearest city and predicted income of the nearest town/city from the preliminary instrumenting regressions reported in Table A.1 and A.2. We also include two additional instruments which are the interactions of predicted travel time and predicted urban income

[15] Among other explanatory variables, household size has significant negative influence on rice sales and land allocated to non-cereal crops particularly during dry season, confirming the importance of subsistence consideration in land allocation and sales. A number of variables representing ownership holding size, and its quality are also statistically relevant in a number of regressions. For instance, an increase in the share of khet, which is particularly suitable for rice production, increases rice sales.

with a dummy for the Terai region. The interaction might capture any regional variations in the effects of the predicted income and travel time as Terai consists of the plain while most of the urban centers are located in the hilly region.[16] Compared with the results in the upper panel of Table 2, there are some changes in the results with respect to the impact of travel time to nearest town/city and its income. The travel time variable has expected negative sign in all regressions except for Herfindahl index of cropland use. It is statistically significant in the equation for share of land devoted to non-cereal crops at 1 percent level and only marginally significant (at 10 percent level) in the case sales of non-rice crops. The travel time variable has a positive sign but is not statistically significant at 10 percent level in the regression for Herfindahl index. Market size of the nearest town/city has expected positive sign and statistical significance only in the regression for share of land devoted to non-cereal crops. The comparison of results reported in the upper and lower panels of Table 2 underlines the importance of taking proper account of the possible endogeneity of road placement as well as urban incomes.

The results in Table 2 suggest that the conventional gravity model does not perform well in identifying the effect of market size on agricultural specialization. Except for non-subsistence specialization index (share of cropland used in non-cereal crop production), the regression results indicate no significant influence of access to nearest urban center or its income on agricultural specialization and commercialization. This may not be entirely unexpected in view of the fact that the nearest town/city is not necessarily the market that a farm household trade with most intensively. In the following section, we report the results when the definition of effective market size is broadened to include urban areas beyond nearest urban town/city.

4.2 The Extent of the Market, Crop Specialization and Commercialization

The weight function which assigns weights to income of urban centers located at different distances play a critical role in defining the effective market size faced by the farm households in a village. There is, however, little guidance from the available empirical literature as to

[16] We emphasize here that the central results of the paper remain intact if we use only the predicted income and travel time as the just identifying instruments.

what a suitable weight function should look like. We rely mainly on the traditional gravity formulation to define the weights. We then check the sensitivity of our results using alternative formulation of the weight function. Our first definition of the effective market size correspond to that given in equation (4) where we assume that $K = 34$.

$$M_i^{eA} = \frac{1}{34}\sum_{j=1}^{34}\frac{m_{ij}}{d_{ij}^2} = \frac{1}{34}\sum_{j=1}^{34} g_{ij}$$

In order to compute M_i^{eA} for village i, we first estimate the gravity measures (urban income$_{ij}$/distance$^2_{ij}$) for all 34 cities using the city income and distance of the city from a village i. Then M_i^{eA} is defined as the average of these gravity measures. We utilize the same procedure to define the predicted instrument (M_i^{eAP}) for M_i^{eA}, where predicted city income and predicted distances are generated from the preliminary regressions discussed earlier. We also include an additional instrument which is defined as the interaction of M_i^{eAP} and a dummy which takes the value of unity if the village is located in Terai (the plains in Nepal) and zero otherwise. Using this definition of the effective market size, equation (5) is estimated by instrumental variable technique. The full regression results along with the diagnostics tests are reported in appendix Table A.3. Appendix Table A.4 reports results from several tests performed to determine the relevance and exogeneity of the instruments. As indicated by the Shea's Partial $\mathrm{R}^2(= 0.59)$ and F-tests of joint significance of the instruments in first stage regression, the instruments explain considerable variations in the extent of market variable. The weak instrument test based on the Cragg-Donald F statistic also resoundingly rejects the null hypothesis that instruments are weak instruments. The overidentification tests also confirm the validity of the instruments used in the IV regressions; the largest value of Hansen's J-statistics is 1.57 with a P-value of 0.21.

The uppermost panel (A) of Table 3 reports the results regarding the effect of the extent of the market. According to Panel A in Table 3, the extent of the market has now statistically significant effect in nearly all regressions, contrary to the findings based on the standard gravity model (Table 2). The coefficient of the extent of market is positive and statistically significant (P-value=0.01 or less) in the regression for percentage of land allocated to non-cereal crops and for sales of crops (non-rice and all). In terms of magnitude, market size has much larger impact on non-rice crop sales (coefficient=0.03) compared with all crops sales including rice (coefficient=0.02). The effect of market size in the case of concentration of crop land use measured by Herfindahl index is, however, negative and statistically significant at 10 percent level. The negative sign of the coefficient implies an increase in the diversity of crop land use with an increase in the extent of market.

In order to tests whether our results regarding the effect of the extent of the market is sensitive to assumption about the weight function, we re-estimate equation (5) using an alternative definition of the weight function. We assume that effect of distance declines at an exponential rate, giving us the following formula for defining the effective market size:

$$M_i^{eB} = \frac{1}{34}\sum_{j=1}^{34} m_{ij} e^{-d_{ij}}$$

The estimates from the IV regressions are reported in Panel B of Table 3. The results indicate that the negative effect of the extent market on Herfindahl index is now statistically significant at 5 percent level. Except for this one difference, the results are nearly the same as in Panel A. We also repeated the analysis defining the effective market size including only 5 nearest cities ($K = 5$) instead of all of Nepalese cities. The overall results are nearly unchanged. Additional robustness checks indicate that our results are not sensitive to the specifications of the weight function and to alternative values of K as long as K is not too small.

The results from the parametric regressions suggest statistically significant and positive effect of the extent of market on percentage of crop devoted to non-cereal crops and sales of non-rice crops and all crops. The results, on the other hand, suggest weak and negative effect of the extent of market on the level of specialization. The results in the case of Herfindahl index of cropland use imply that larger market size induces increasing diversity in land use at the village level. This seems to be counter-intuitive and contrary to the Smithian conjecture about the role of markets in fostering greater division of labor. As already noted, price and production risks are among the most important factors influencing farm households cropping and market participation decisions. Because of the critical influence of price risks, the relationship between the extent of market and agricultural specialization is likely to be non-linear. In the following section, we explore this possibility using semi-parametric technique.

4.3 Non-Linearity and Stages of Specialization

A semi-parametric specification of equation (6) can be used as the basis for exploring the non-linearity in the relationship between market size and different measures of agricultural specialization. Equation (6) assumes that the relationships between dependent variable y_i^z and explanatory variables in vector X'_{zi} takes parametric form. We utilized the semi-parametric estimation technique proposed by Robinson (1988) and Yatchew(1998) to estimate the function $g_z(.)$ which describes the relationship between y_i^z (measures of specialization and commercialization) and M_i^{eA} (market size). The estimator involves stepwise procedure to estimate $g_z(.)$. At the first step, y_i^z and all explanatory variables in vector X'_{zi} are purged off the effect of M_i^{eA} using standard non-parametric kernel regressions. Next, the residuals generated from the kernel regressions are then used to estimate the coefficient vector $\overset{\wedge}{\gamma}_z$. The effects of explanatory variables in vector X'_{zi} are then taken out of y_i^z using estimated coefficient vector $\overset{\wedge}{\gamma}_z$. Finally, a standard kernel regression is run with residual of y_i^z from the preceding step as the dependent variable and M_i^{eA} as the explanatory variable. This final kernel regression provides the estimate of $g_z(.)$. We follow this stepwise procedure with one modification. To correct for the possible endogeneity of M_i^{eA}, we use the standard control function approach and include residual from the first stage regression as an additional explanatory variable in the semi-parametric regression.

The estimated $g_z(.)$ functions along with its 95% confidence intervals are plotted in Figure 3 and 4. According to Figure 3, the relationship between market size and Herfindahl index of concentration of cropland use is U-shaped. The estimated function has a tight 95% confidence interval for much of the range of urban market size except for very high values. The share of land allocated to non-cereal crops increases with an increase in the level of urban income, and the relationship appears to be nearly linear. In the case of sales (Figure 4), both sales of all crops and of non-rice crops increases with an increase in the urban market size. The relationship has a steeper slope in the case of non-rice sales compared with all crop sales. A comparison of the two curves shows that for the entire range of urban income, the curve for non-rice sales lies above that for all crops including rice. This is consistent with the Engel curve prediction that non-rice crops face a higher income elasticity relative to rice, a subsistence crop. Figure 4 also points to slight flattening of the curves at high level of urban market size. However, confidence interval at this high levels of urban market size is very large raising concerns about the statistical significance of this portion of the curve.

Consistent with the case of Herfindahl index, the 95% confidence interval is wide at higher values of the extent of market in all of the other graphs. This prompts us to check the statistical significance of the non-linearity found especially in the case of Herfindahl index. To this end, we re-estimate all regressions by adding a square of the extent of the market term in the parametric specification in equation (5). This squared term is instrumented using the square of the predicted extent of the market variable. We also include interaction of this term with a dummy for Terai region as an additional instrument. The IV diagnostics show that the instrument set is acceptable both in terms of relevance and exogeneity criteria. The IV regression results are reported in Table 4. In the regression for Herfindahl index, the level and square of the extent of the market variable have now become highly statistically significant at 1 percent level. The signs of these two terms also confirm that the relationship between the extent of the market and agricultural specialization is U shaped. In the case of percentage of land devoted to non-cereal crops, the squared term is not statistically significant. This confirms the semi-parametric results that the relationship between land allocated to non-cereal crops and the extent of the market is nearly linear. In the sales regressions, addition of squared term renders both level and squared terms individually statistically insignificant at 5 percent level perhaps due to multicollinearity. Comparison with the results in Table 3 shows that the relationship between sales and the extent of the market can be more suitably described as linear.We also checked robustness of our results by repeating the semi-parametric and parametric analysis in the case where the extent of the market is measured by M_i^{eB}. The results are nearly identical.

The convex relationship between the Herfindahl index of cropland use and market size may appear puzzling at first sight. However, if this curve is compared with that for share of land area planted with non-cereal crops, the reason becomes apparent. The higher values of Herfindahl index at lower levels of urban market size is due to the fact that the households in villages with access to smaller markets are basically self-sufficient and thus specialize in cereal crops. As market size faced by a village increases, it starts producing wider range of crops resulting in lower value of the Herfindahl index. However, as market size crosses a threshold, an increase in market size encourages more specialization with villages specializing more and more in non-cereal crops. This is consistent with the findings reported in Imbs and Wacziarg(2003). With access to only small markets, subsistence considerations predominate and land is used primarily to produce subsistence cereal crops. This is corroborated by the fact that there is little commercialization of agriculture when villages have limited access to largest markets. When the village has access to larger markets, diversified consumer demand in the urban areas induce farmers to allocate more land to high value non-cereal crops. As market size increases further, it ensures more stable trading opportunity and a reduction in the price uncertainty for the non-subsistence crops thereby allowing farmers to specialize in non-subsistence crops (e.g. fruits and vegetables).

5 Conclusions

The process of structural change that transforms a traditional subsistence based self-sufficient village economy into a more market oriented and specialized one is an important part of the long-run evolution of an economy (Locay, 1990; Gollin et. al., 2002). This transformation process is of great importance to a majority of the developing countries where agriculture- still the mainstay of economic activity- is characterized by low levels of commercialization and specialization. Policy makers in these countries grapple with ways to accelerate the transformation of agriculture. The objective of this paper is to analyze structural transformation within the agricultural sector, especially as it relates to the extent of market.

Using village level data from Nepal, we analyze two dimensions of structural change in agriculture: the pattern of product diversification (crop specialization), and the degree of market participation (i.e., commercialization). The pattern of crop specialization is measured by the Herfindahl index of concentration of cropland use, and by the share of land devoted to non-cereal (non-rice) crops. Sales of non-rice crops and all crops as a percentage of total production are taken as measures of commercialization. As opposed to the standard practice of defining market as the nearest urban center, we introduce a broader measure of the extent of market which incorporates the possibility that villagers may trade at multiple urban locations. The empirical analysis also corrects for the possible endogeneity of the extent of the market.

The parametric regression results indicatea statistically significant and positive effect of the extent of market on the share of land devoted to non-cereal crops and sales of non-rice and all crops. The effect of market size is negative, albeit statistically significant only at 10 percent level, in the case of Herfindahl index. The semi-parametric regression analysis suggests that the relationship between the extent of market and Herfindahl index of specialization is U-shaped. No such non-linearity is uncovered in the case of crop sales and share of land devoted to non-cereal crops.The results imply that when the farmers do not have access to large urban markets, crop production is dominated by subsistence considerations with villages specializing in the production of staple cereal crops. With an increase in the extent of market, crop production first becomes diversified with farmers producing both cereal and non-cereal crops. As the extent of market crosses a critical threshold, villages begin to specialize again– this time in the production of non-cereal crops. The evidence on the stages of agricultural diversification is thus similar to the stages identified earlier in the literature for the economy as a whole and for the manufacturing sector (Imbs and Wacziarg, 2003).

The paper demonstrates the importance of market size in inducing agricultural specialization and commercialization. This evidence confirms the crucial role of improving farmers' access to markets, through investment in transport infrastructure and removal of barriers to trading both at the domestic and international level.

References

1. Ades, Alberto and Edward Glaeser, 1999, "Evidence on Growth, Increasing Returns and the Extent of the Market," *Quarterly Journal of Economics*, Vol.., p. 1025-1045.

2. Alcala, Francisco, and Antonio Ciccone, 2003, "Trade, Extent of the Market, and Economic Growth 1960-1996," *Working paper.*

3. Buera, F and J. kaboski (2006): The Rise of the Service Economy, Working Paper, Northwestern University and Ohio State University.

4. Binswanger, H, and Klaus Deininger, 1997, "Explaining Agricultural and Agrarian Policies in Developing Countries," *Journal of Economic Literature*, Vol. 35(4), p.1958-2005.

5. Chenery, H, Robinson, M. Surquin (1986): Industrialization and Growth: A Comparative Study, Oxford University Press.

6. Ciccone, A and K. Matsuyama (1996): Start-up Costs and Pecuniary Externalities as barriers to Economic Development, Journal of Development Economics, April, 1996.

7. Dixit, A (1993): "Scale Economies, Technological Change and Diversification", in Braverman, Hoff, and Stiglitz eds: The Economics of Rural Organization, Oxford University Press.

8. Fafchamps, M. and F. Shilpi. 2003. "The Spatial Division of Labor in Nepal." *Journal of Development Studies*, Vol. 39, p. 23–66.

9. Fafchamps, Marcel, and Forhad Shilpi, 2005. "Cities and Specialization: Evidence from South Asia." *Economic Journal*,Vol. 115, April.

10. Frankel, J and D. Romer (1999): Does Trade Cause Growth?, *American Economic Review*, June, 1999.

11. Gollin, Douglas, Stephen Parente and Richard Rogerson, 2002, "The Role of Agriculture in Development," *American Economic Review Papers amd Proceedings*, Vol. 92(2), p.160-164.

12. Gollin, Douglas, Stephen Parente and Richard Rogerson, 2006, "The Food Problem and the Evolution of International Income Levels," forthcoming in *Journal of Monetary Ecnomics*.

13. Imbs, Jean and Romain Wacziarg, 2003, "Stages of Diversification," *American Economic Review*, Vol. 93(1), p.63-86.

14. Islam, N and S. Thomas (1996): Foodgrain Price Stabilization in Developing Countries, Food Policy Review 3, IFPRI, Washington DC.

15. Jacoby, Hanan G., 2000, "Access to Markets and the Benefits of Rural Roads," *Economic Journal*, Vol.110, p.713-737.

16. Johnson, D.G, 2000, "Population, Food and Knowledge", *American Economic Review*, Vol. 90(1), p.1-14.

17. Johnston, Bruce and John Mellor, 1995, "The Role of Agriculture in Economic Development", in G. Peters ed., *Agricultural Economics*, Elgar Publishers, Brookfield VT.

18. Kalemli-Ozcan, sebnem, Brent E. Sorensen, and Oved Yosha, 2003, "Risk Sharing and Industrial Specialization: Regional and International Evidence," *American Economic Review*, Vol. 93(3), p.903-918.

19. Kuznets, S (1973): Modern Economic Growth: Findings and Reflections (Nobel Lecture), *American Economic Review*, 63, 247-258.

20. Laitner, J, 2000: "Structural Change and Economic Growth", *Review of Economic Studies*, July, 2000.

21. Lewis, A (1954): "Economic Development with Unilimited Supplies of Labor", *Manchester School*, 22 (2), 139-191.

22. Locay, L (1990): "Economic Development and the Division of Production Between Households and Markets", *Journal of Political Economy*, October, 1990.

23. Lucas, R (1988): "The Mecahnics of Economic Development", Journal of Monetary Economics.

24. Lucas, R (2004): "Life Earnings and Rural-urban migration", Journal of Policitical Economy, February, 2004.

25. Matsuyama, K (2005): Structural Change, Working Paper, NorthWestern University.

26. Melitz, M and G. Ottaviano, 2005, "Market Size, Trade, and Productivity," *NBER Working Paper #11393*.

27. Murphy, Kevin, Andrew Shleifer, and Robert Vishny, 1989, "Industrialization and the Big Push," *Journal of Political Economy,* Vol. 97(5), p.1003-1026.

28. Newbery, D and J. Stiglitz (1981): *The Theory of Commodity Price Stabilization*, Oxford University Press.

29. Pingali, Prabhu, 2004, "Agricultural Diversification: opportunities and Constraints," Food and Agriculture Organization, *mimeo.*

30. Robinson, P. M., 1988, "Root-n-Consistent Semiprarametric Regression," *Econometrica*, Vol. 56(4), p.931-954.

31. Rodriguez-Clare, A (1996): "The Divison of Labor and Economic Development", *Journal of Development Economics*, June 1996.

32. Rosenstein-Rodan, Paul, 1943, "Problems of Industrialization of East and South-Eastern Europe, *Economic Journal*, Vol.53, p.202-211.

33. Rose grant, M., and M. Sombilla, 1997, "Critical Issues suggested by trends in Food, Population and the Environment to the Year 2020," *American Journal of Agricultural Economics*, Vol. 79(5), p.1497-70.

34. Roumasset, J (1976): *Rice and Risk*, North Holland.

35. Schultz, T.W., 1968, *Economic Growth and Agriculture*, New York:McGraw-Hill.

36. Smith, Adam, 1976, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Chicago: The University of Chicago Press.

37. Syverson, C, 2004, "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, Vol. 112(6), p.1181-1222.

38. von Thünen, J. H. 1966, *Isolated State: an English Edition of Der Isolierte Staat*, Oxford: Pergamon Press.

39. Wooldridge, J (2007): Econometric Analysis of Cross Section and Panel Data, second Edition, MIT Press.

40. World Bank, 1982, *World Development Report*, New York: Oxford University Press.

41. Yatchew, Adonis, 1998, "Nonparametric Regression Techniques in Ecnomics, *Journal of Economic Literature*, Vol. 36(2), p.669-721.

Table 1: Agricultural Specialization and Characteristics of Nearest Urban Market

	Median	Mean	SD	Min	Max	No. of observation
A. Specialization Indices						
Herfindahl index of concentration of cropland use	0.272	0.308	0.129	0.136	1	235
Share of total cultivated land devoted to non-cereal crops	0.251	0.317	0.251	0	1	235
Percentage of production sold						
Rice	0.104	0.146	0.16	0	0.7	231
Non-Rice	0.149	0.206	0.185	0	0.882	237
All crops	0.151	0.19	0.155	0	0.724	237
B. Market Size and Access						
Distance to nearest town/city (hours)	2.35	4.43	4.98	0.047	29.55	237
Total income in nearest town/city (Rs. Million)	301.8	1209.91	2592.8	73	10748	237
Effective market size	16.81	104.98	573.02	0.196	7957.8	237

Table 2: Gravity model and Agricultural Sepcialization

	Without Endogeneity Correction			
	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice	All crop
Travel time to nearest town/city (log)	0.012	-0.082	-0.033	-0.030
	(0.64)	(4.17)***	(2.47)**	(2.88)***
Total Income of nearest town/city (log)	-0.002	0.031	-0.001	-0.002
	(0.21)	(2.17)**	(0.07)	(0.28)
Observations	235	235	237	237
R-squared	0.19	0.35	0.36	0.43
	With Endogeneity Correction			
Travel time to nearest town/city (log)	0.026	-0.112	-0.030	-0.019
	(1.62)	(4.69)***	(2.02)**	(1.60)
Total Income of nearest town/city (log)	-0.009	0.054	0.003	0.004
	(0.89)	(2.63)***	(0.19)	(0.36)
Observations	235	235	237	237
R-squared	0.18	0.34	0.36	0.42

Note: Regression controls are same as in appendix Table A.3

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Market Size and Agricultural Specialization

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
		Panel A		
Log(Effective market size)	-0.017	0.097	0.033	0.023
-Defintion A	(1.68)*	(5.18)***	(3.06)***	(2.70)***
R-squared	0.18	0.33	0.38	0.44
Over Identification Test				
Hansen's J Statistics	0.001	1.57	0.19	1.12
P-value	0.97	0.21	0.66	0.29
		Panel B		
Log(Effective market size)	-0.021	0.084	0.028	0.021
-Definition B	(2.08)**	(4.29)***	(2.87)***	(2.74)***
R-squared	0.19	0.3	0.38	0.45
Over Identification Test				
Hansen's J Statistics	0.007	0.37	0.27	0.76
P-value	0.94	0.54	0.6	0.38

Definition A: Effective market size= Σ_i(urban income$_i$/distance$_i^2$) where i=1…34 cities/towns in Nepal

Definition B: Effective market size= Σ_i(urban income$_i$/e$^{\text{distancei}}$) where i=1…34 cities/towns in Nepal

Note: Results from instrumental variable estimation. Regression controls are same as in appendix Table A.3

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Non-Linear Impact of Market Size on Agricultural Specialization

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
Log(Effective market size)	-0.137	0.165	-0.035	-0.017
-Defintion A	(4.08)***	(2.43)**	(0.98)	(0.57)
Log(Effective market size)^2	0.018	-0.010	0.010	0.006
	(3.55)***	(0.99)	(1.81)*	(1.34)
R-squared	0.28	0.35	0.36	0.43
Over Identification Test				
Hansen's J Statistics	0.74	1.94	0.62	1.25
P-value	0.69	0.38	0.73	0.54

Definition A: Effective market size= Σ_i(urban income$_i$/distance$_i^2$) where i=1…34 cities/towns in Nepal

Note: Results from instrumental variable estimation. Regression controls are same as in appendix Table A.3

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.1: Instrumenting equation for urban income
Dependent Variable: log(Total urban income)

	Unit	Coefficient
Area of district in which town/city is located	log(square	-0.591
		(1.43)
Distance to nearest navigable part of river	log(km)	-0.672
		(1.11)
Arable land area of the district	log(ha)	-0.244
		(1.02)
Mean elevation	meters	0.000
		(0.08)
Standard deviation of district elevation	meters	0.002
		(0.89)
Mountain Region	yes=1	-0.706
		(0.89)
Intercept		31.584
		(4.50)**
Number of observations		34
R-squared		0.43

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table A.2. Instrumenting equation for travel time
Dependent Variable: log(Travel time between village and town/city)

	Unit	Coefficient	T-statistics
Foot Travel time between village and town/city	log	0.677	174.120
Village Characteristics			
Area of district in which village is located	log(square	0.202	21.650
Total areable land of district	log(ha)	-0.145	-12.820
Mean elevation	meters	0.000	14.020
Standard deviation of district elevation	meters	0.000	-6.710
Central region	yes=1	-0.229	-24.060
West region	yes=1	-0.281	-26.650
Mid-west region	yes=1	-0.219	-14.900
Far-west region	yes=1	-0.199	-11.980
Town/City Characteristics			
Area of district in which town/city is located	log(square	0.067	5.340
Distance to nearest navigable part of river	log(km)	0.048	7.280
Arable land area of the district	log(ha)	-0.005	-0.300
Mean elevation	meters	0.000	8.180
Standard deviation of district elevation	meters	0.000	-6.350
Central region	yes=1	-0.150	-11.400
West region	yes=1	-0.186	-15.860
Mid-west region	yes=1	-0.172	-9.550
Far-west region	yes=1	-0.082	-4.200
Intercept		-1.780	-9.370
Number of observations		9187	
R-squared		0.8443	

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table A.3: Agricultural Specialization and the Extent of Market: Full regressions

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	Sales as % of production All Crops
Log(Effective market size)[1]	-0.017	0.097	0.033	0.023
	(1.68)*	(5.18)***	(3.06)***	(2.70)***
Household Size (log)	0.037	-0.428	-0.070	-0.106
	(0.50)	(3.66)***	(0.81)	(1.54)
Share of adult female	-0.047	-0.968	0.228	0.209
	(0.25)	(3.08)***	(1.13)	(1.20)
Share of children	-0.167	0.196	0.008	0.001
	(0.86)	(0.69)	(0.04)	(0.01)
Share of Old	0.273	-0.139	-0.475	-0.103
	(1.35)	(0.37)	(1.86)*	(0.53)
Average female education (log)	0.130	0.414	-0.146	-0.054
	(1.45)	(3.56)***	(1.38)	(0.69)
Average male education (log)	-0.116	-0.072	0.031	0.084
	(2.24)**	(0.83)	(0.60)	(1.96)*
Average land area owned (log)	-0.032	0.251	0.078	0.132
	(0.77)	(3.23)***	(1.63)	(3.15)***
Sahre of irrigated land (owned)	-0.051	0.068	0.022	0.076
	(1.61)	(1.15)	(0.43)	(1.84)*
Share of Khet owned	-0.035	-0.073	0.015	-0.062
	(0.86)	(0.82)	(0.23)	(1.25)
Average value of farm equipments (log)	-0.003	-0.033	0.009	0.008
	(0.40)	(2.38)**	(0.98)	(1.13)
Share of owned land of quality Awal	0.099	-0.125	0.082	0.105
	(1.50)	(1.12)	(1.13)	(1.76)*
Share of owned land of quality Dwaim	0.034	-0.093	0.081	0.061
	(0.73)	(1.03)	(1.35)	(1.18)
Share of owned land of quality Sim	-0.028	0.096	-0.057	-0.058
	(0.49)	(1.00)	(1.33)	(1.46)
Number of farm animals (log)	-0.038	0.056	-0.033	-0.024
	(1.59)	(1.39)	(1.08)	(0.99)
Per capita consumption expend. (log)	0.008	0.087	0.022	-0.007
	(0.33)	(1.79)*	(0.78)	(0.31)
If there is bank in the village	0.010	-0.132	-0.025	0.000
	(0.49)	(3.64)***	(0.95)	(0.00)
Remittance income (log)	0.003	-0.013	0.002	0.001
	(1.56)	(3.19)***	(0.57)	(0.46)
Ecological belt dummy (Mountain=1)	0.067	-0.034	-0.095	-0.104
	(1.73)*	(0.51)	(2.12)**	(2.51)**
Ecological belt dummy (Hill=1)	0.072	-0.008	-0.081	-0.089
	(2.71)***	(0.17)	(1.99)**	(2.47)**
Intercept	0.270	0.253	-0.060	0.273
	(1.19)	(0.58)	(0.22)	(1.21)
R-squared	0.18	0.33	0.38	0.44
Observations	235	235	237	237

Note: Results from instrumental variable estimation.

1/:Definition: Effective market size= Σ_i(urban income$_i$/distance$_i^2$) where i=1…34 cities and towns in Nepal

Robust t statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.4 :IV regression Diagonistics

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
Relevance instruments in 1st stage regression				
Shea's Partial R^2	0.59	0.59	0.59	0.59
F-Statistics	130.2	130.2	131.8	131.8
P-value	0	0	0	0
Weak Identification Test				
Cragg-Donald F statistic	152.7	152.7	156.2	156.2
Stock-Yogo weak ID test Critical Value				
10% Maximum IV size	19.93	19.93	19.93	19.93
Over Identification Test				
Hansen's J Statistics	0.001	1.57	0.19	1.12
P-value	0.97	0.21	0.66	0.29

Figure 1: Location of Markets

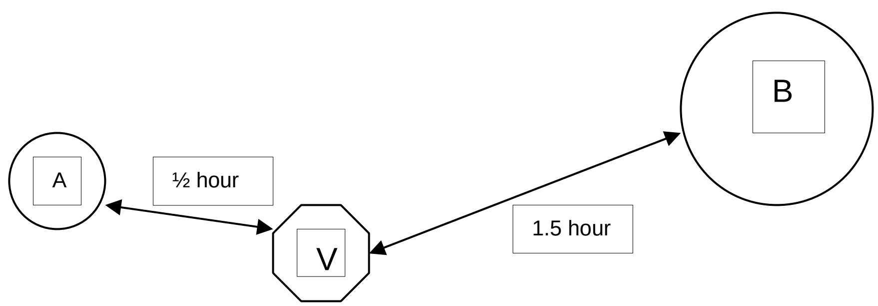

Figure 2a. Travel time to Nearest Town/City

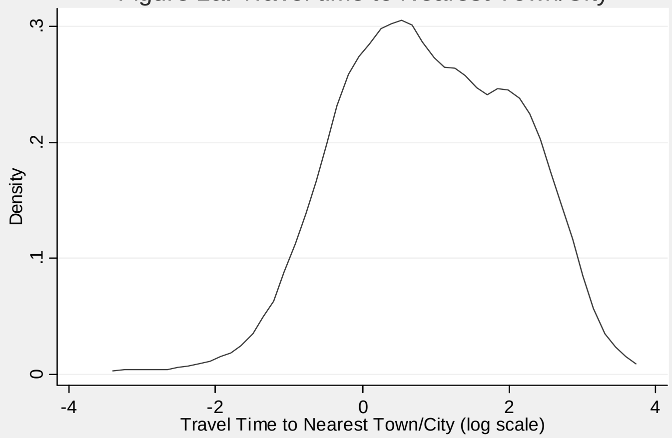

Figure 2b. Distribution of Total Urban Income

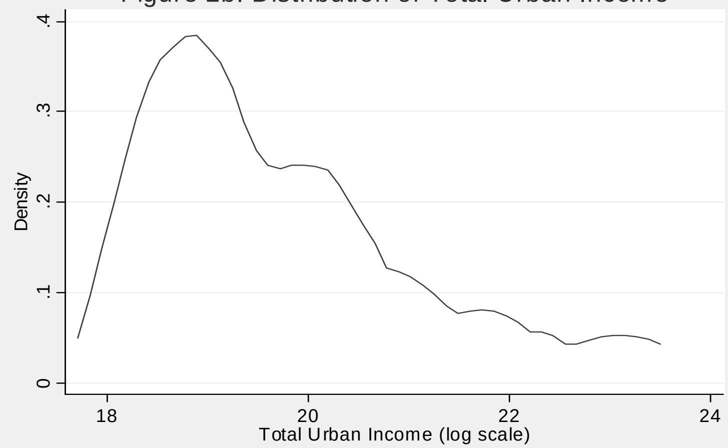

Figure 2c. Distribution of Extent of Market

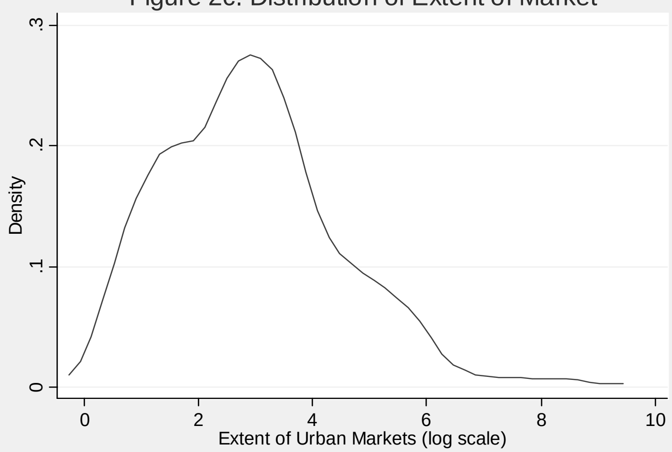

Figure 3. Agricultural Specialization and the extent of the market

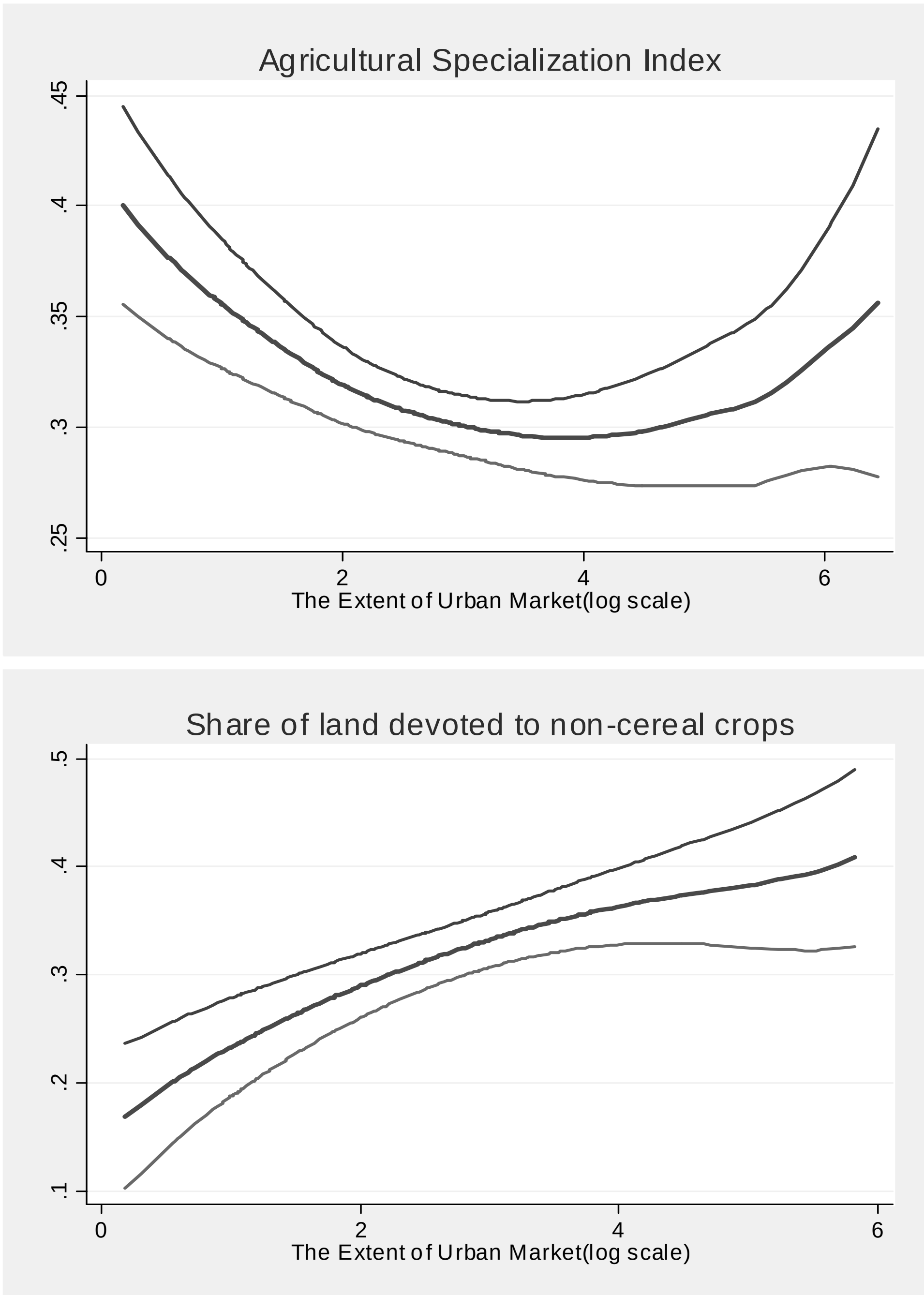

Figure 4. Agricultural commercialization and the extent of the market